

Money grows trees: a socio-ecological path analysis



Jason Walker
Arizona State University
Global Institute for Sustainability
urbanecologist@gmail.com



Abstract

This study analyzed the interplay among socioeconomics, patterns of urban development, and their effect on vegetation dynamics in the Phoenix metropolitan area. I hypothesized that income is the driving factor of vegetation coverage, primarily affecting neighborhood characteristics, which in turn have the largest influence on vegetation. I developed a conceptual model for the relationships among income, neighborhood characteristics and vegetation abundance. This model (and alternative models) were tested via path analysis, a statistical evaluation of the hypothesized path model and the empirical data. This analytical procedure allowed the direct and indirect effects of economic and housing variables to be incorporated and tested within a single hypothetical model, which explicitly accounts for the interdependence of dependent variables. Results from the path analysis indicate income, neighborhood age and housing density have significant effects on tree coverage, as well as significant relationships among those variables. Analysis of the direct versus indirect effects on tree coverage indicate that income predominately drives tree coverage directly, and not indirectly through neighborhood characteristics as hypothesized.

What is path analysis and why use it

Path analysis is a more general form of multiple regression, which tests the predetermined theoretical effects of dependent variables on each other and their direct and indirect effects on an independent variable. This procedure provides means for statistical control of interdependency among the data within a theoretical framework. Controlling such interdependency is important in nearly all socio-ecological systems due to the large degree of collinearity among social and ecological phenomena. In order to test a theoretical framework, an *a priori* model (or more commonly a series of such models) must be generated from a theoretical framework, as seen below.

Theoretical Framework

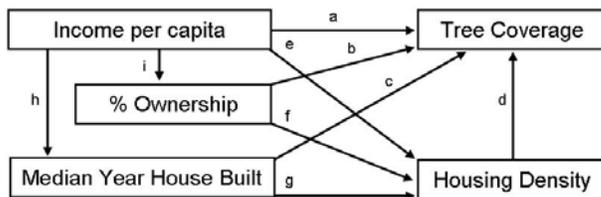


Figure 2. Baseline Model.

Tree Coverage in residential neighborhoods is directly affected by:

- Neighborhood Age (+), measured as Median Year House Built (-)
The Phoenix metropolis is a young urban development. Assuming trees live 50+ years, logic holds that biomass should covary with neighborhood age in young urban developments, representing secondary succession in an urban environment.
- Ownership (+), measured as Percent Houses Owned (+)
Greater ownership of single detached homes should enhance the abundance of vegetation due to a greater interest in maintaining property values. However, this effect may be confounded by vegetation enhancement of apartment complexes.
- Income (+), measured as Income Per Capita (+)
Economics have already been shown to affect plant biodiversity within the Phoenix metropolitan (Hope et al. 2003). I predict this effect will translate to vegetation coverage as well.
- Urban Sprawl (-), measured as Housing Density (+)
The proportion of impervious surface (e.g. buildings and roads) constrains the available growing space for vegetation to grow. Thus, vegetation coverage should negatively vary with increasing impervious surface coverage.

Housing Density is directly affected by:

- Income (+), measured as Income Per Capita (+)
Richer neighborhoods will consist of denser houses, largely due to the indirect effect of neighborhood age on housing density.
- Ownership (-), measured as Percent Houses Owned (+)
Renters typically occupy high density apartment complexes, and owners occupy detached houses.
- Neighborhood Age (+), measured as Median Year House Built (-)
As Phoenix continues to grow, space has become limited and therefore more expensive, causing developers to more densely now than in the decades before present.

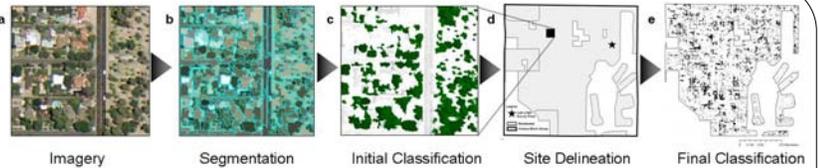
Neighborhood Age is directly affected by:

- Income (+), measured as Income Per Capita (+)
Richer people prefer newer neighborhoods.

Ownership is directly affected by:

- Income (+), measured as Income Per Capita (+)
Richer people have a higher probability of owning a house.

Urban Forest Classification



Sampling Strategy

CAP LTER conducts a semi-decadal ecological survey at 206 sites selected via a dual-density, stratified-random tessellation design, which samples the Phoenix metropolitan area's urban core with triple the spatial intensity as the surrounding Sonoran Desert. Of those sites, 112 were identified as within a census-defined block group containing residential land use as defined by the Matricos Association of Governments (MAG). Sampling sites for the urban forest classification were delineated from these three geographical datasets; in which, a single sample was identified as the area classified as residential by MAG within a census block group containing a CAP LTER survey point (see d above). This approach was conducted in order to limit the analysis of vegetation coverage to residential neighborhoods, corresponding to the same sampling extent the US Census Bureau gathered data on socioeconomics and housing variables.

Remote Sensing

Following site delineation, we employed an object-oriented approach to classify urban vegetation through a hybrid of image segmentation and rule-based classification. To more accurately estimate real world objects, the image was apportioned into basic units for analysis at the object-appropriate scale before classification can occur through a process of segmentation (b). Segmentation was conducted based on contextual information (i.e. within-pixel spectra values and patch texture) as well as neighborhood characteristics making possible the extraction of real-world objects, proper in shape, as the basic units for analysis. Following segmentation, the objects were subjected to an urban forest classification scheme (c) developed for high-resolution (0.61m), true-color (red, green, blue), aerial photography. The outset of this procedure produces a binary matrix where the entire raster set is classified highlighting the elements of the urban forest (e) for the specified areas (d). To determine the accuracy of the classification, an extensive groundtruthing campaign was conducted. Subsequent analysis of commission errors indicated a user's accuracy of woody vegetation of 0.96, indicating that 96% of the objects identified as woody vegetation were, indeed, trees or shrubs.

Path Analysis

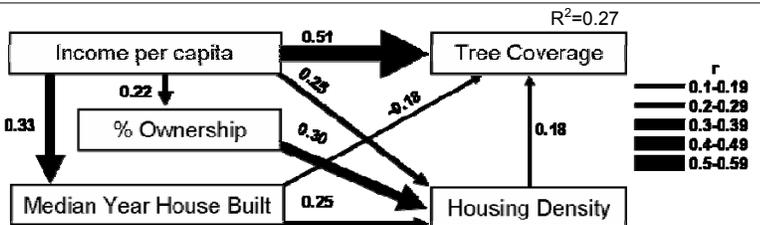


Figure 2. Final model. $\chi^2=3.89$, $df=2$, p -value=0.143, $RMSEA=0.093$. All paths are significant at $\alpha=0.05$

Model Selection

Model	χ^2	df	p-value	RMSEA
Baseline Model	3.89	2	0.073	0.093
Model a	31.1	2	<0.001	0.204
Model b	3.09	2	0.213	0.085
Model c	4.09	2	0.133	0.146
Model d	3.09	2	0.213	0.146
Candidate: Income path b				
Model: Model A	3.09	2	0.213	0.085
Model e	10.9	3	0.133	0.190
Model f	10.9	3	0.089	0.182
Model g	10.6	3	0.076	0.191
Candidate: Neighborhood age				

Table 1. Path analysis decision tree. Analysis of significance in path analysis differs from regression in that one is interested in a model that does not contradict the data, or statistically speaking, does not reject the null hypothesis. Backwards pruning was conducted by removing paths from baseline model (depicted in Figure 1) and assessing the improvement of the model. Deletion of path b from the base model was warranted suggested by an enhancement of the overall model's p -value, suggesting that there is not a clear cut effect of ownership on vegetation abundance. All other paths were retained.

Direct vs. Indirect Effects

	Income			Year Built			Density			% Ownership		
	Tot	Dir	Ind	Tot	Dir	Ind	Tot	Dir	Ind	Tot	Dir	Ind
Tree Coverage	0.50	0.51	0.01	0.05	--	0.05	0.18	0.18	--	0.05	--	0.05
% Ownership	0.22	--	--	--	--	--	--	--	--	--	--	--
Median Year	0.33	--	--	--	--	--	--	--	--	--	--	--
Density	0.26	0.24	0.02	0.30	0.30	--	--	--	--	0.30	0.30	--

Table 2. Total (Tot), Direct (Dir), and Indirect (Ind) Effects of the final model. Direct effects represent correlations coefficients between variables while indirect effects represent the effect of one variable on another through its relationship with intermediary variables. This analysis suggests that income's effect on vegetation coverage through its relationship to neighborhood characteristics is minimal, suggesting the relationship below is likely the most appropriate.

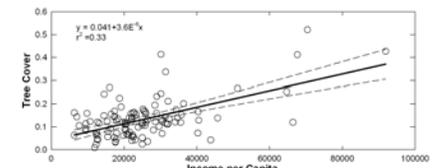


Figure 3. Socioeconomics drive biomass too. Simple regression with untransformed variables. Solid line represents the estimated regression line, whereas the dashed lines represent the 95% confidence interval.

Acknowledgements

This project was directed by the Central Arizona-Phoenix Long Term Ecological Research Project (CAP LTER). Additional support was provided by the Integrative Graduate Education and Research Training (IGERT) in Urban Ecology at Arizona State University.

Conclusions

Richer People have:
- More trees ✓
- Newer neighborhoods ✓
- Denser neighborhoods ✓
- Higher Ownership ✓

Newer Neighborhoods have:
- Fewer Trees ✓
- Higher Density ✓

Denser Neighborhoods have:
- More Trees X

Higher Ownership Rates have:
- Higher Density ✓

Housing variables appear to have a minimal (yet significant) affect on vegetation coverage. It appears income fundamentally (and directly) drives vegetation abundance, rather than neighborhood characteristics.